RELATIONSHIPS BETWEEN SPECIAL CLASSES OF MATRIX TRANSFORMATIONS

A Thesis

Presented to

the Department of Mathematics

Emporia Kansas State College

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

by

Linda Ingold Campbell

August 1974

George D. Poole
Approved for the Major Department

Approved for the Graduate Council

ACKNOWLEDGMENTS

I extend my deep appreciation to Dr. George Poole for his assistance and encouragement and to all my professors who made this study possible.

Acknowledgment and appreciation are also expressed to my husband and to my parents for their many words of encouragement.

TABLE OF CONTENTS

Chapter I

INTRODUCTION

While most texts in the area of matrix theory describe and classify various types of matrices, many of these texts fail to develop the relationships between specific classes of matrices.

The objective of this thesis is to compare certain classes of matrices. These matrices are compared according to characteristics inherent to particular classes of matrices and also according to the geometrical interpretations that certain matrices represent.

When writing in matrix theory, one must assume that the reader has a basic knowledge of matrices. However, Chapter II contains basic and essential definitions of matrix theory as well as a clarification of the notation used in this thesis.

Chapter III contains the definitions of the matrices that are discussed in this thesis with examples to illustrate the structures of these matrices.

Chapter IV contains a comparison of the eigenvalues among the various classifications of matrices.

Chapter V contains some geometrical properties of three transformations while Chapter VI contains a more detailed study of one of these transformations.

Chapter II

BASIC DEFINITIONS AND NOTATION

The notation used in matrix theory varies from text to text. This chapter contains the notation used in this thesis as well as the definitions basic to matrix theory.

DEFINITION 2.1. A system or ordered sextuplet $V = (V, F, +, \cdot, \oplus, \odot)$ is called a <u>vector space over the field</u> F if and only if:

a) $(F, +, \cdot)$ is a field F whose identity elements are denoted by θ and 1;

b) $(V, \oplus)$ is an abelian group whose identity element is denoted by 0;

c) For all $\alpha, \beta \in F$ and $v, w \in V$ where $\alpha \odot v \in V$, the following are true:

i) $\alpha \odot (v \oplus w) = (\alpha \odot v) \oplus (\alpha \odot w)$
ii) $(\alpha + \beta) \odot v = (\alpha \odot v) \oplus (\beta \odot v)$
iii) $\alpha \odot (\beta \odot v) = (\alpha \cdot \beta) \odot v$
iv) $1 \odot v = v$

The above notation, while precise, is cumbersome. It is conventional to adopt the following abbreviated notation:

i) $\alpha(v + w) = \alpha v + \alpha w$
ii) $(\alpha + \beta) v = \alpha v + \beta v$
iii) $\alpha(\beta v) = (\alpha\beta) v$
iv) $1v = v$

The following notation will be used consistently in this paper:

a) F will be a field.

b) Lower case Greek letters will be elements of F. These elements will be called scalars in this paper.

c) Capital Latin letters will denote vector spaces over F.

d) Lower case Latin letters will denote elements of vector spaces. These elements will be called <u>vectors</u>.

DEFINITION 2.2. If V is a vector space and if $S = (v_1, \ldots, v_n) \subseteq V$, then the set S is linearly <u>dependent</u> over F if there exist elements $\lambda_1, \ldots, \lambda_n$ in F, not all of them 0, such that $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$. If the set S is not linearly dependent over F, then it is said to be linearly <u>independent</u> over F.

DEFINITION 2.3. A maximal linearly independent subset of a vector space V is called a <u>basis</u> of V.

NOTE: In this paper, only vector spaces containing finite maximal linearly independent sets are considered. Furthermore, it can be shown that all bases for a vector space contain the same number of elements.

DEFINITION 2.4. The <u>dimension</u> n of a vector space is the number of elements in any basis of V over F.

DEFINITION 2.5. A linear transformation T from a vector space V to a vector space W, both over the scalar field F, is a mapping of V into W such that for all v, w $\in$ V and for all $\alpha, \beta \in F$, $(\alpha v + \beta w)T = \alpha(vT) + \beta(wT)$.

Let V_m be a vector space with an arbitrary but fixed basis $(v_1, \ldots, v_m)$. Let T be a linear transformation of V_m into a vector space W_n and let $(w_1, \ldots, w_n)$ be any fixed basis for W_n. For each $i = 1, 2, \ldots, m$ v_iT is a uniquely determined vector of W_n and hence is uniquely represented as a linear combination of the w_j, $j = 1, 2, \ldots, n$:

$$\begin{aligned} v_1T &= \alpha_{11}w_1 + \alpha_{12}w_2 + \cdots + \alpha_{1n}w_n \\ v_2T &= \alpha_{21}w_1 + \alpha_{22}w_2 + \cdots + \alpha_{2n}w_n \\ &\vdots \\ v_mT &= \alpha_{m1}w_1 + \alpha_{m2}w_2 + \cdots + \alpha_{mn}w_n \end{aligned}$$

Notice the meaning of the subscripts. The first subscript i of α_{ij} means that α_{ij} is one of the coefficients of the representative of the vector v_iT relative to the w-basis, and the second subscript j of α_{ij} means that α_{ij} is the coefficient of w_j in that representation. Relative to the two bases, T is completely determined by the mn subscripts of the coefficient α_{ij} together with this interpretation of the meaning of the coefficients. It is necessary to pay attention to the order of the basis vectors to avoid ambiguity. With this convention understood, T can be represented by the rectangular array of scalars,

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

DEFINITION 2.6. A rectangular array containing m rows and n columns of elements of a field F is called an <u>m x n matrix</u> over F.

DEFINITION 2.7. The <u>sum</u> of two matrices is the matrix of the sum of the two corresponding linear mappings.

DEFINITION 2.8. If A is an m x n matrix and B is a p x m matrix, then the <u>product</u> BA is the matrix of the composite of the linear mappings corresponding to B and A.

All matrices used in this paper are assumed to be square matrices over the complex number field ($F = \mathbb{C}$).

NOTATION USED IN THIS THESIS

A^T	Transpose of A
A^*	Conjugate transpose of A
$\mathbb{C}$	Complex number field
∎	End of proof
I	Identity matrix
A^{-1}	Inverse of A
$\|A\|$	Determinant of A
$[a_{ij}]$	The matrix with entries a_{ij}
V_n	Denotes a vector space of dimension over an appropiate scalar field

Chapter III

COMPARISON OF MATRICES BY STRUCTURE

The following classes of matrices were chosen because of their interesting relationships to one another.

DEFINITION 3.1. A matrix A is normal if and only if $AA^* = A^*A$.

$$\begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix} \text{ is normal.}$$

DEFINITION 3.2. A matrix A is unitary if and only if $A^{-1} = A^*$.

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ is unitary.}$$

DEFINITION 3.3. A matrix A is Hermitian if and only if $A = A^*$.

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ is Hermitian.}$$

DEFINITION 3.4. A matrix $A = (a_{ij})$ is diagonal if and only if $a_{ij} = 0$ whenever $i \neq j$.

$$\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \text{ is diagonal.}$$

DEFINITION 3.5. A matrix A is Gramian if and only if for some B, $A = BB^*$.

$$\begin{pmatrix} 2 & i \\ -i & 5 \end{pmatrix} \text{ is Gramian; } B = \begin{pmatrix} 1 & 1 \\ i & -2i \end{pmatrix}.$$

DEFINITION 3.6. A matrix A is <u>idempotent</u> if and only if $A^2 = A$.

$$\begin{pmatrix} 4 & -2 \\ 6 & -3 \end{pmatrix} \text{ is idempotent.}$$

DEFINITION 3.7. A matrix A is <u>involutory</u> if and only if $A^2 = I$.

$$\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \text{ is involutory.}$$

DEFINITION 3.8. A matrix A is <u>symmetric</u> if and only if $A = A^T$.

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ is symmetric.}$$

DEFINITION 3.9. A matrix A is <u>orthogonal</u> if and only if $A^T = A^{-1}$.

$$\begin{pmatrix} i & -\sqrt{2} \\ \sqrt{2} & i \end{pmatrix} \text{is orthogonal.}$$

DEFINITION 3.10. A matrix A is <u>semi-normal</u> if and only if $AA^T = A^TA$.

$$\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \text{ is semi-normal.}$$

Before Definition 3.11 can be given, it is necessary to give a short explanation concerning the similarity of two matrices. The matrix A is similar to the matrix B if there exists a nonsingular matrix P such that $A = PBP^{-1}$. Geometrically, A is similar to B if A and B are the same transformations but represented by different bases.

DEFINITION 3.11. A matrix A is <u>simple</u> if and only if A is similar to a diagonal matrix.

$\begin{pmatrix} 1 & 2 \\ 8 & 3 \end{pmatrix}$ is simple; $P = \begin{pmatrix} 1 & 1 \\ 4 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$.

The next part of this chapter contains theorems which establish relationships between particular classes of the above defined matrices. The last part of this chapter contains examples of matrices to illustrate these relationships with a diagram at the end of the chapter which shows explicitly the exact connections between the types of matrices.

THEOREM 3.1. (a) Unitary matrices are normal.
(b) Hermitian matrices are normal.
(c) Diagonal matrices are normal.

Proof: To establish (a), assume $A^* = A^{-1}$ and, hence, $AA^* = AA^{-1}$. Therefore, $AA^* = I$. Similarly, $A^*A = I$. Therefore $AA^* = A^*A$.

To establish (b), assume $A = A^*$. Therefore, $AA^* = A^*A^* = A^*A$.

To establish (c), assume $A = [a_{ij}]$ where $a_{ij} = 0$ whenever $i \neq j$. Hence $[a_{ij}]^* = [\overline{a_{ij}}]$. Therefore $[a_{ij}]\,[a_{ij}]^* = [a_{ij}]\,[\overline{a_{ij}}] = [\overline{a_{ij}}][a_{ij}] = [a_{ij}]^*\,[a_{ij}]$. ∎

THEOREM 3.2. A real unitary matrix is real orthogonal.

Proof: Assume $A^{-1} = A^*$. Since A is real, $A^* = A^T$. Therefore $A^{-1} = A^T$. ∎

THEOREM 3.3. A real Hermitian matrix is symmetric and a real symmetric matrix is Hermitian.

Proof: Assume $A = A^*$. Since A is real $A^* = A^T$. Therefore, $A = A^T$. The second statement of the theorem follows similarly by assuming $A = A^T$ first. ∎

THEOREM 3.4. (a) If A is unitary and Hermitian, then A is involutory.
(b) If A is unitary and involutory, then A is Hermitian.
(c) If A is Hermitian and involutory, then A is unitary.

Proof: To prove (a); assume $A^* = A^{-1}$ and $A = A^*$. It follows that $A = A^{-1}$. The proofs of (b) and (c) are similar. ∎

THEOREM 3.5. A Gramian matrix is Hermitian.

Proof: Assume $A = BB^*$. Then, $A^* = (BB^*)^* = (B^*)^*B^* = BB^*$. Therefore, $A = A^*$. ∎

THEOREM 3.6. A real Gramian matrix is symmetric.

Proof: Assume $A = BB^*$. By Theorem 3.5, $A = A^*$. Since A is real, $A^* = A^T$. Therefore $A = A^T$. ∎

THEOREM 3.7. Any diagonal matrix is symmetric.

Proof: The proof is clear. ∎

THEOREM 3.8. A symmetric matrix is semi-normal.

Proof: Assume $A = A^T$. Then $AA^T = A^TA^T = A^TA$. ∎

THEOREM 3.9. A normal matrix with real elements is semi-normal.

Proof: The proof is clear. ∎

THEOREM 3.10. (a) If a matrix is both involutory and symmetric, then it is orthogonal.
(b) If a matrix is involutory and orthogonal, then it is symmetric.
(c) If a matrix is symmetric and orthogonal, then it is involutory.

Proof: To establish (a), assume $A = A^{-1}$ and $A = A^T$. Then it follows that $A^{-1} = A^T$. The proofs for (b) and (c) are similar. ∎

THEOREM 3.11. A matrix that is both idempotent and unitary must be the identity.

Proof: Assume $A^2 = A$ and $AA^* = I$. Then $A = A \cdot I = A(AA^*) = (AA)A^* = AA^* = I$. ∎

THEOREM 3.12. Normal matrices are simple.

Proof: Assume $AA^* = A^*A$. By the Schur triangularization theorem (page 67, [3]), $T = UAU^*$, where T is upper triangular and U is unitary. So $TT^* = UAU^*UA^*U^* = UAA^*U^*$. Also, $T^*T = UA^*U^*UAU^* = UA^*AU$. Since $AA^* = A^*A$, it follows that $UAA^*U^* = UA^*AU^*$ and therefore $TT^* = T^*T$. This implies that T is diagonal. ∎

THEOREM 3.13. Idempotent matrices are simple.

Proof: Assume $A^2 = A$. A is similar to a Jordan form; that is, $A = PJP^{-1}$. Since $A^2 = A$, this means that $PJP^{-1}PJP^{-1} = PJP^{-1}$, so PJ^2P^{-1} and $J^2 = J$. The Jordan form has submatrices along the diagonal which take the form of, for example, either

$$\begin{bmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{bmatrix} \text{ or } \begin{bmatrix} \lambda_i & 0 & 0 \\ 0 & \lambda_i & 0 \\ 0 & 0 & \lambda_i \end{bmatrix}.$$

Since λ is either 0 or 1 (Theorem 4.5), these matrices become $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, or $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

If all the diagonal blocks take the form of either one of the last two matrices, then J is diagonal. If one diagonal block takes the form of either of the first two matrices, then $J^2 \neq J$. Therefore J must be diagonal. ∎

THEOREM 3.14. A matrix that is both idempotent and normal is Hermitian.

Proof: Since A is normal, by Theorem 3.12 $A = UDU^*$ where U is unitary and D is diagonal. Hence since $A^2 = A$, then $UDU^*UDU^* = UDU^*$; so $UD^2U^* = UDU^*$ which means that $D^2 = D$. Therefore each $d_{ij} = 1$ or 0. Also $A^* = UD^*U^*$. Therefore $D = D^*$ and $A^* = UDU^* = A$. ∎

Not only do most texts in matrix theory fail to give examples, they do not always illustrate the relationships among the various kinds of matrices. In giving examples for the various relationships, it becomes necessary to show that while certain classes of matrices are related these classes are not equal. In fact, in many cases particular classes are proper subsets of other classes, or in some cases are even the intersection between classes. This next section contains examples of the matrices which were defined at the beginning of this chapter, along with examples of relationships which have been stated in the theorems in this chapter.

Example 3.1. (a) A matrix that is simple but not normal is

$$\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}; \; P = \begin{pmatrix} 1 & 1 \\ 4 & -2 \end{pmatrix}, \; B = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}.$$

(b) A matrix that is simple, but not semi-normal is

$$\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}.$$

(c) A matrix that is simple but not idempotent is

$$\begin{pmatrix} 1+i & 1 \\ 1 & 1+i \end{pmatrix}; \; P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \; D = \begin{pmatrix} i+2 & 0 \\ 0 & i \end{pmatrix}.$$

Example 3.2. (a) A matrix that is normal but not semi-normal is

$$\begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}.$$

(b) A matrix that is semi-normal but not normal is

$$\begin{pmatrix} 1 & i \\ i & i \end{pmatrix}.$$

(c) A matrix that is both semi-normal and normal is

$$\begin{pmatrix} 3 & 2 \\ 2 & -1 \end{pmatrix}.$$

Example 3.3. (a) A matrix that is idempotent but not normal is

$$\begin{pmatrix} 4 & -2 \\ 6 & -3 \end{pmatrix}.$$

(b) A matrix that is normal but not idempotent is

$$\begin{pmatrix} 1+i & 1 \\ 1 & 1+i \end{pmatrix}.$$

(c) A matrix that is both normal and idempotent (Hermitian) is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Example 3.4. A matrix that is normal but not Hermitian, unitary, or orthogonal is

$$\begin{pmatrix} 1+i & 1 \\ 1 & 1+i \end{pmatrix}.$$

Example 3.5. (a) A matrix that is unitary but not orthogonal is

$$\begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}.$$

(b) A matrix that is orthogonal but not unitary is

$$\begin{pmatrix} \frac{i}{\sqrt{2}} & -\sqrt{2} \\ \sqrt{2} & i \end{pmatrix}.$$

(c) A matrix that is both unitary and orthogonal (real orthogonal) is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Example 3.6. (a) A matrix that is symmetric but not unitary is

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

(b) A matrix that is unitary but not symmetric is

$$\begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}.$$

(c) A matrix that is both unitary and symmetric is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Example 3.7. (a) A matrix that is unitary but not Hermitian is

$$\begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}.$$

(b) A matrix that is Hermitian but not unitary is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

(c) A matrix that is both Hermitian and unitary (involutory) is

$$\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Example 3.8. (a) A Gramian matrix that is neither idempotent nor unitary is

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix};\ B = \begin{pmatrix} i & -i \\ 1+i & 1 \end{pmatrix}.$$

(b) A Gramian matrix that is idempotent is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix};\ B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

(c) A Hermitian matrix that is not Gramian is

$$\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Note: To show that a matrix is not Gramian, assume that it is. Thus $A = BB^*$ means that the entries in the identical positions in the two matrices must be equal. If equating these elements results either in an inconsistent system of equations or in some other mathematical contradiction, then the matrix is not Gramian.

Example 3.9. (a) A matrix that is semi-normal but not symmetric is

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

(b) A matrix that is both semi-normal and symmetric is

$$\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Example 3.10. (a) A matrix that is symmetric but not Hermitian is

$$\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}.$$

(b) A matrix that is Hermitian but not symmetric is

$$\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

(c) A matrix that is both Hermitian and symmetric (real) is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Example 3.11. (a) A matrix that is symmetric but not diagonal is

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

(b) A matrix that is both symmetric and diagonal is

$$\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Example 3.12. (a) A matrix that is Gramian but not symmetric is

$$\begin{pmatrix} 2 & i \\ -i & 5 \end{pmatrix}: \; B = \begin{pmatrix} 1 & 1 \\ i & -2i \end{pmatrix}.$$

(b) A matrix that is symmetric but not Gramian is

$$\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}.$$

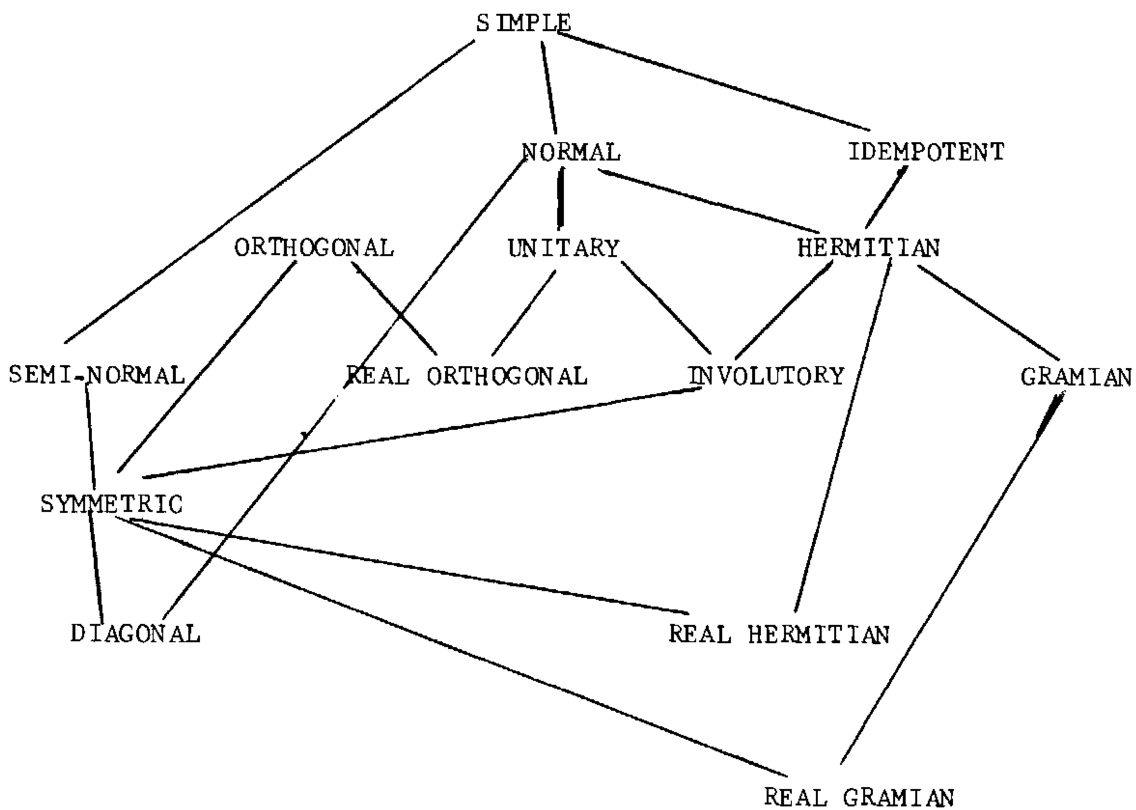
(c) A matrix that is both symmetric and Gramian is

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}; \; B = \begin{pmatrix} i & -i \\ 1+i & 1 \end{pmatrix}.$$

On the last page of this chapter, there is a diagram which shows the relationships which have been established in the theorems and illustrated by examples. An explanation of the diagram is as follows. When there is a single line between two classes of matrices, with one class above the other, the class that is listed above contains the class from which the line is leading upward. For example, there are single lines leading to simple from semi-normal, normal, and idempotent. These lines indicate that the sets of all semi-normal, normal, and idempotent matrices are subsets of the set of all simple matrices. When there are two lines from two different sets of matrices leading down to a single set, the single set is the intersection of the other two sets. For examplé, there is a line leading to idempotent and one leading to normal both of which lead from Hermitian. This means that when idempotent classes and normal classes intersect, the class that is formed is Hermitian.

A DIAGRAM SHOWING THE RELATIONSHIP

OF CLASSES OF MATRICES



Chapter IV

EIGENVALUES AND EIGENVECTORS

This chapter contains a discussion of eigenvalues and eigenvectors. It is necessary to first define the terms "eigenvalue" and "eigenvector." After these terms have been defined, it becomes particularly useful to give specific examples of the geometric interpretations of various eigenvectors. Once the terms have been defined and the geometrical interpretation has been given, it is possible to show the relationships between the defined classes of matrices by using the sets of eigenvalues which belong to each class.

DEFINITION 4.1. (a) An eigenvalue of a matrix A is a scalar λ such that $Ax = \lambda x$ for some vector $x \neq 0$.

(b) The vector x is called an eigenvector of the matrix A.

DISCUSSION: If λ is an eigenvalue of A, then $Ax - \lambda x = 0$ and $(A - \lambda I)x = 0$. Thus λ is a scalar such that the above homogeneous set of equations has a nontrivial solution; that is, a solution other than $x = 0$. Thus an eigenvalue is a scalar λ such that $|A - \lambda I| = 0$.

Many books describe the method for finding eigenvalues and eigenvectors, but few books contain any additional explanation as to their value. A fev examples can show some of the uses for eigenvalues and eigenvectors. Before giving some specific problems, it is necessary to give a geometric interpretation of eigenvectors.

Let $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$. The eigenvalues of A are $\lambda_1 = 3$ and $\lambda_2 = 2$. Every eigenvector associated with λ_1 is of the form $(k, 0)^T$, where k is any nonzero scalar, since $\begin{pmatrix} 3-3 & 0 \\ 0 & 2-3 \end{pmatrix}\begin{pmatrix} k \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Furthermore, the set of vectors of the form $(k, 0)^T$ is such that $A(k, 0)^T = \lambda_1(k, 0)^T$: that is, $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} k \\ 0 \end{pmatrix} = 3\begin{pmatrix} k \\ 0 \end{pmatrix}$.

Hence the set of eigenvectors belonging to $\lambda_1 = 3$ is mapped onto itself under the transformation represented by A, and the image of each eigenvector is a fixed scalar multiple of that eigenvector. This fixed scalar is precisely the eigenvalue with which the set of eigenvectors is associated.

Similarly, every vector associated with λ_2 is of the form $(0, k)^T$ where k is any nonzero scalar. The set of vectors of the form $(0, k)^T$ is such that $A(0, k)^T = \lambda_2(0, k)^T$; that is, $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ k \end{pmatrix} = 2\begin{pmatrix} 0 \\ k \end{pmatrix}$.

Hence, the set of eigenvectors associated with $\lambda_2 = 2$ is mapped onto itself under the transformation represented by A, and the image of each eigenvector is a fixed scalar multiple of the eigenvalue. The fixed scalar multiple is λ_2; that is, 2.

The sets of vectors of the forms $(k, 0)^T$ and $(0, k)^T$ lie along the x-axis and y-axis respectively. (See Figure 4.1)

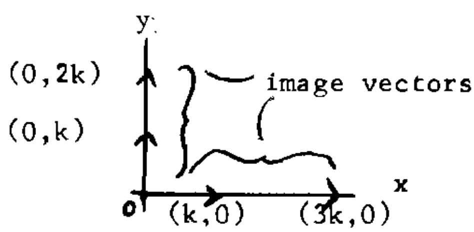


Figure 4.1

Under a magnification of the plane represented by the matrix $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, the one-dimensional vector spaces containing the sets of vectors of the forms $(k, 0)^T$ and $(0, k)^T$ are mapped onto themselves, respectively, and are called <u>invariant</u> <u>vector</u> <u>spaces</u>.

Problem 4.1. Determine the invariant vector spaces under a shear parallel to the x-axis represented by the matrix A, where $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$.

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 1$. Associated with each eigenvalue is the set of eigenvectors of the form $(k, 0)^T$, where k is any nonzero scalar. Then $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} k \\ 0 \end{pmatrix} = 1\begin{pmatrix} k \\ 0 \end{pmatrix}$ and the one-dimensional vector space containing the set of vectors of the form $(k, 0)^T$ is an invariant vector space. Furthermore, since $\lambda_1 = \lambda_2 = 1$, each vector in the vector space is its own image under A.

Problem 4.2. Determine the invariant vector space under a projection of the plane represented by the matrix $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$.

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 0$. Associated with the eigenvalue $\lambda_1 = 1$ is the set of eigenvectors of the form $(k, k)^T$ where k is any nonzero scalar; associated with the eigenvalue $\lambda_2 = 0$ is the set of eigenvectors of the form $(0, k)^T$ where k is any nonzero scalar. Then, $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} k \\ k \end{pmatrix} = 1\begin{pmatrix} k \\ k \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ k \end{pmatrix} = 0\begin{pmatrix} 0 \\ k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Since the vectors of the form $(0, k)^T$ are mapped onto the zero vector, the one-dimensional vector space containing these vectors is mapped <u>into</u> but not <u>onto</u> itself and the space is <u>not</u> considered an invariant vector space. However, the one-dimensional vector space containing the set of vectors of the form $(k, k)^T$ is an invariant vector space. Note that these vectors lie along the line $y = x$ and that the plane is mapped onto this line under the projection of the plane represented by the matrix A.

Problem 4.3. Discuss the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

To find the eigenvalues, $\begin{vmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} = 0$. Thus $(1-\lambda)^2 - 1 = 0$. So, $1-2\lambda+\lambda^2-1 = 0$. Therefore $\lambda(\lambda-2) = 0$ and $\lambda_1 = 0$ and $\lambda_2 = 2$.

Case 1. To find the eigenvectors for $\lambda_1 = 0$, let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Then $Ax = \lambda_1 x$ is written $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda_1\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Thus, $x_1 - x_2 = 0$ and $-x_1 + x_2 = 0$ which means that $x_1 = x_2$. The solution is the set of all $[x_1, x_1]$ that is, the one-dimensional subspace spanned by $[1, 1]$. Note that these eigenvectors are all the nonzero scalar multiples of $[1, 1]$. These eigenvectors comprise the kernel of the mapping, except that 0 is in the kernel, but is not an eigenvector.

Case 2. To find the eigenvectors for $\lambda_2 = 2$, let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Then, $Ax = \lambda_2 x$ means that $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. The solution set is the set of all $[x_1, -x_1]$; that is, the one-dimensional subspaces spanned by $[1, -1]$.

Particular sets of eigenvalues are associated with certain classes of matrices. These eigenvalues are related in the same manner as are the types of matrices. The following theorems deal with the sets of eigenvalues which belong to the defined classes of matrices.

THEOREM 4.1. The eigenvalues of a Hermitian matrix are real numbers.

Proof: Assume $A = A^*$. Let λ be an eigenvalue of A, so $Ax = \lambda x$; $x \neq 0$. Hence, $x^*Ax = \lambda x^*x$. Then, $(x^*Ax)^* = (\lambda x^*x)^*$. It follows that $x^*A^*x = \bar{\lambda}x^*x$. Therefore $x^*A^*x = x^*Ax = \lambda x^*x = \bar{\lambda}x^*x$ which means that $\lambda = \bar{\lambda}$; that is, λ is real. ∎

THEOREM 4.2. The eigenvalues of a unitary matrix are (real or) complex numbers of Modulus (absolute value) 1.

Proof: Assume $A^{-1} = A^*$ (also expressed as $AA^* = I$). Let λ be a root of A so that $Ax = \lambda x$; $x \neq 0$. Then, $(Ax)^* = (\lambda x)^*$ which means that $x^*A^* = \bar{\lambda}x^*$. Hence, $x^*A^*Ax = \bar{\lambda}x^*\lambda x = \bar{\lambda}\lambda x^*x$. Thus, $x^*Ix = \bar{\lambda}\lambda x^*x = x^*x$. Therefore $\bar{\lambda}\lambda = 1$ so $|\lambda| = 1$. ∎

THEOREM 4.3. The eigenvalues of a real orthogonal matrix are either 1 or -1.

Proof: Assume $A^{-1} = A^T$ (that is $A^TA = I$). Let λ be a root of A so that $Ax = \lambda x$. Then, $(Ax)^T = (\lambda x)^T$ so that $x^TA^T = \lambda x^T$. Hence, $x^TA^TAx = \lambda x^T \lambda x$. So, $x^TIx = \lambda\lambda x^Tx = x^Tx$. Therefore, $\lambda\lambda = 1$ which means that $\lambda^2 = 1$ and $\lambda = \pm 1$. ∎

THEOREM 4.4. The eigenvalues of an involutory matrix are either -1 or 1.

Proof: Assume $A^2 = I$. Let λ be a root of A so that $Ax = \lambda x$; $x \neq 0$. Hence, $AAx = A\lambda x = \lambda Ax = \lambda\lambda x$. Hence, $Ix = \lambda\lambda x = x$ so that $\lambda\lambda = 1$. Therefore $\lambda^2 = 1$, and $\lambda = \pm 1$. ∎

THEOREM 4.5. The eigenvalues of an idempotent matrix are either 0 or 1.

Proof: Assume $A^2 = A$. Let λ be a root of A so that $Ax = \lambda x$; $x \neq 0$. Then, $AAx = A^2x = \lambda Ax = \lambda\lambda x$. So, $Ax = \lambda\lambda x$ which means that $\lambda x = \lambda\lambda x$. Therefore, $0 = \lambda^2 x - \lambda x = (\lambda^2 - \lambda)x$. It follows that $\lambda^2 - \lambda = 0$ since $x \neq 0$. So, $\lambda(\lambda - 1) = 0$ and $\lambda = 0$ or $\lambda = 1$. ∎

THEOREM 4.6. The eigenvalues of a Gramian matrix are real.

Proof: The proof is the same as the proof for Theorem 4.1. ∎

THEOREM 4.7. The eigenvalues of a diagonal matrix A may be any complex number.

Proof: Assume A is diagonal. Then the eigenvalues for A are equal to the entries along the diagonal of A. Since these entries may be either real, purely imaginary or complex, it follows that the eigenvalues will be the same. ∎

THEOREM 4.8. The eigenvalues of a normal matrix may be any complex number.

Proof: Theorem 3.1 (c) states that all diagonal matrices are normal. Therefore, the set of eigenvalues for diagonal matrices must be included in the set of eigenvalues for normal matrices. Theorem 4.7 states that these eigenvalues may be any complex number. ∎

THEOREM 4.9. The eigenvalues of a symmetric matrix may be any complex number.

Proof: Theorem 3.7 states that all diagonal matrices are symmetric. This proof is then similar to the proof of Theorem 4.8. ∎

THEOREM 4.10. The eigenvalues of a simple matrix may be any complex number.

Proof: Theorem 3.13 states that all normal matrices are simple. This means that the set of eigenvalues of normal matrices must be included in the set of eigenvalues for simple matrices. This set of eigenvalues, as stated in Theorem 4.8, is real, purely imaginary, or complex. ∎

Chapter V

SOME GEOMETRICAL INTERPRETATIONS OF THREE TRANSFORMATIONS

Since matrices represent transformations, it would be remiss to omit some geometrical consequences of some transformations in a paper which deals with matrices. To describe all transformations would be too ambitious; so to give the reader a general idea of transformations, three specific transformations have been selected. The remainder of this chapter contains some interesting geometrical interpretations of orthogonal, unitary, and similarity transformations.

DEFINITION 5.1. If U is an orthogonal matrix, then the transformation $Y = UX$ is called an orthogonal transformation.

THEOREM 5.1. The determinant of an orthogonal matrix is 1 or -1.

Proof: Assume $U^T = U^{-1}$. Then $UU^T = I$. Hence $|U|\,|U^T| = |U|^2 = 1$. Therefore, $|U| = \pm 1$. ∎

DEFINITION 5.2. If U is an orthogonal matrix and $|U| = +1$, the transformation $Y = UX$ is proper orthogonal. If $|U| = -1$, then the transformation is improper orthogonal.

DEFINITION 5.3. If U is a unitary matrix, then the transformation $Y = UX$ is a unitary transformation.

DEFINITION 5.4. The formation of a matrix $U = B^{-1}AB$ is a <u>similarity</u> transformation of A.

THEOREM 5.2. An orthogonal transformation preserves distance.

Proof: Let $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ be the Cartesian coordinates $(x_1, x_2, \ldots, x_n)$ of a point P in an Euclidean n-dimensional space. Then, $X^TX = x_1^2 + x_2^2 + \ldots + x_n^2$ gives the square of the distance the origin to P.

Let $Y = AX$ (A is an n^{th} order matrix) where $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$. The distance of Y from the origin is $Y^TY = y_1^2 + y_2^2 + \ldots + y_n^2$. Hence $Y^TY = (AX)^TAX = X^TA^TAX$. If $A^TA = I$, then $Y^TY = X^TX$; that is, distance is preserved. To say that $A^TA = I$ is to say that A is orthogonal. Therefore $Y = AX$ is an orthogonal transformation that preserves distance.∎

THEOREM 5.3. An orthogonal transformation leaves the angle between any two vectors unchanged.

Proof: Assume X and X' are two vectors in n-dimensional space. Let $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ be the coordinates of P and $X' = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}$ be the coordinates of Q (see Figure 5.1).

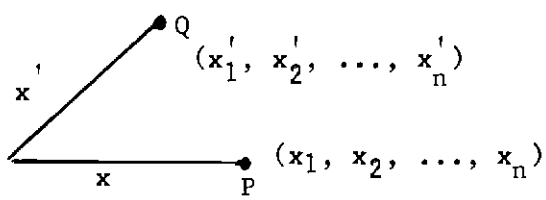


Figure 5.1

The angle between X and X' is defined as: $X \cdot X' = |X|\ |X'|\cos\theta$.

Thus, $\cos\theta = \dfrac{x_1x_1' + x_2x_2' + \ldots + x_nx_n'}{(x_1^2 + x_2^2 + \ldots + x_n^2)^{\frac{1}{2}}\left[(x_1')^2 + (x_2')^2 + \ldots + (x_n')^2\right]^{\frac{1}{2}}}$

In matrix form this becomes $\dfrac{X^TX'}{(X^TX)^{\frac{1}{2}}\ (X'^TX')^{\frac{1}{2}}}$. Let X be transformed into Y by an orthogonal trensformation Y = AX, and X' transformed into Y' so that Y' = AX'. Then ϕ is the angle between Y and Y'. Similarly,

$$\cos\phi = \frac{Y^TY'}{(Y^TY)^{\frac{1}{2}}\ (Y'^TY')^{\frac{1}{2}}} = \frac{(AX)^TAX'}{\left((AX)^TAX\right)^{\frac{1}{2}}\ \left((AX')^TAX'\right)^{\frac{1}{2}}} =$$

$$\frac{X^TA^TAX'}{(X^TA^TAX)^{\frac{1}{2}}\ (X'^TA^TAX')^{\frac{1}{2}}} = \frac{X^TX'}{(X^TX)^{\frac{1}{2}}\ (X'^TX)^{\frac{1}{2}}} = \cos\theta.$$

Thus the angle between the two vectors is invariant. ∎

A SPECIAL RESULT OF THEOREM 5.3.

Consider any rotation of axes in $\mathcal{E}_2$ and let the angle of rotation be θ. Then the unit vectors E_1 and E_2 of the y-coordinate system are the unit vectors $(\cos\theta, \sin\theta)$ and $(-\sin\theta, \cos\theta)$ respectively in the x-coordinate system. (See Figure 5.2)

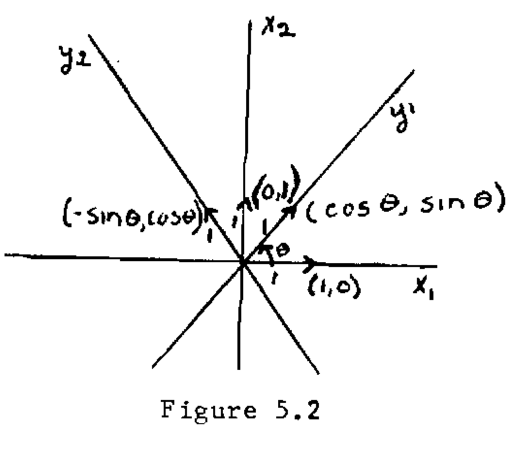


Figure 5.2

Hence the transformation is:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (1)$$

the matrix of which is orthogonal and has determinant 1. Conversely, let

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (2)$$

be any orthogonal transformation in $\mathcal{E}_2$. Here the unit vectors E_1 and E_2 of the y-coordinate system are the orthogonal unit vectors (a_{11}, a_{21}) and (a_{12}, a_{22}) respectively in the x-coordinate system.

Let θ in the transformation of the type (1) be chosen in such a way that the unit vector $(a_{11}, a_{21}) = (\cos\theta, \sin\theta)$ defines the y_1-axis. Then the y_2-axis, being orthogonal to the y_1-axis, is defined either by the unit vector $(-\sin\theta, \cos\theta)$ or by the unit vector $(\sin\theta, -\cos\theta)$ in the opposite direction (see Figure 5.3). That is, either $(a_{12}, a_{22}) = (-\sin\theta, \cos\theta)$ or $(a_{12}, a_{22}) = (\sin\theta, -\cos\theta)$. The transformation as shown in Figure 5.2 is thus either a rotation or a transformation of the form:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (3)$$

which is an improper orthogonal transformation since the determinant is -1. Statement (3) may be written $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

and hence be interpreted as the product of the proper orthogonal transformation $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$ and the transformation $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ which reverses the choice of positive direction on the y_2-axis. This latter transformation is called a <u>reflection in the y_1-axis</u>.

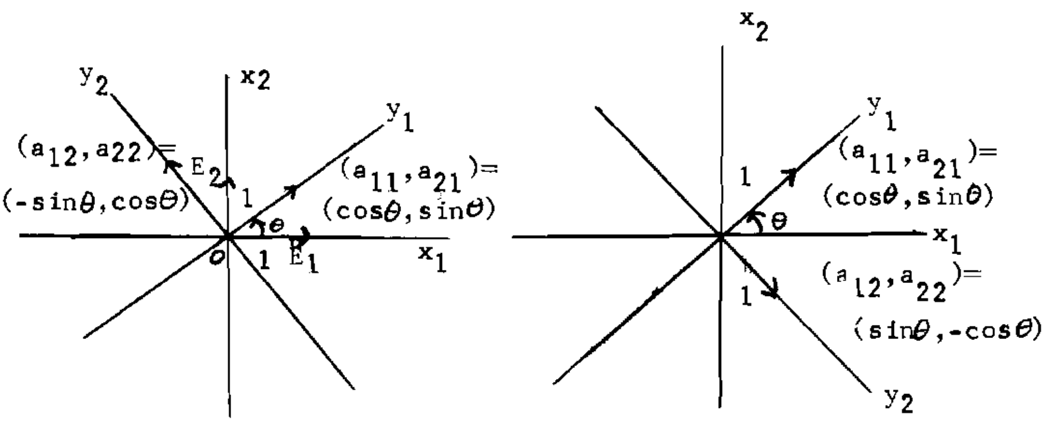


Figure 5.3

This argument may be extended to $\mathcal{E}_3$. That is, in $\mathcal{E}_3$, the proper orthogonal transformations represent rotations of axes and the improper orthogonal transformations represent rotations combined with a reflection in one of the coordinate planes.

THEOREM 5.4. The scalar product of two plane vectors is a scalar invariant under an orthogonal transformation. That is, $\vec{a}' = \vec{a_1}'i + \vec{a_2}'j$ and $\vec{b}' = \vec{b_1}'i + \vec{b_2}'j$ are the image vectors of $\vec{a} = \vec{a_1}i + \vec{a_2}j$ and $\vec{b} = \vec{b_1}i + \vec{b_2}j$ respectively under a rotation of the plane about the origin or under a reflection of the plane with respect to a line through the origin; then, $a_1b_1 + a_2b_2 = a_1'b_1' + a_2'b_2'$.

Proof: Let R be either a rotation matrix or a reflection matrix. Then, $(a_1', a_2')^T = R(a_1, a_2)^T$ and $[(a_1', a_2')^T]^T = [R(a_1, a_2)^T]^T$. This means that $(a_1', a_2') = (a_1, a_2)R^T$. Similarly, $R(b_1, b_2)^T = (b_1', b_2')^T$. Then multiplying equals by equals: $(a_1, a_2)R^TR(b_1, b_2)^T = (a_1', a_2')(b_1', b_2')^T$. Since R is orthogonal, $(a_1, a_2)(b_1, b_2)^T = (a_1', a_2')(b_1', b_2')^T$ and $a_1b_1 + a_2b_2 = a_1'b_1' + a_2'b_2'$. ∎

THEOREM 5.5 An orthogonal transformation preserves the separation of two vectors; that is, the distance between two vectors.

Proof: $|x - y|$ is the separation of x and y. Let A be an orthogonal transformation. Then, $|Ax - Ay|^T |Ax - Ay| = \left[|Ax|^T - |Ay|^T\right]\left[|Ax - Ay|\right] = \left[|x^TA^T - y^TA^T|\right]\left[|Ax - Ay|\right] = |x^TA^TAx - x^TA^TAy - y^TA^TAx + y^TA^TAy| = |x^Tx - x^Ty - y^Tx + y^Ty| = |x^T(x - y) - y^T(x - y)| = |x^T - y^T||x - y| = |x - y|^T|x - y|$. ∎

THEOREM 5.6. A unitary transformation leaves distance invariant.

Proof: Let $X^*X = x_1^*x_1 + x_2^*x_2 + \ldots + x_n^*x_n = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2$. Consider a unitary transformation $Y = AX$ (A is an n^{th} order matrix). Then $Y^*Y = (AX)^*AX) = X^*A^*AX = X^*X$. Therefore distance is preserved when A is a unitary matrix. ∎

THEOREM 5.7. A unitary transformation leaves the angle between two vectors invariant.

Proof: Parts of this proof are similar to the proof of Theorem 5.3, and have been omitted here. Let X' and X be two vectors with θ the angle between them. Then, $\cos\theta = \frac{X^*X'}{(X^*X)^{\frac{1}{2}}(X'^*X')^{\frac{1}{2}}}$. Consider the unitary transformation $Y = AX$ and $Y' = AX'$ with ϕ as the angle between Y and Y'. Therefore, $\cos\phi = \frac{Y^*Y'}{(Y^*Y)^{\frac{1}{2}}(Y'^*Y')^{\frac{1}{2}}} =$

$$\frac{(AX)^*(AX')}{\left[(AX)^*AX\right]^{\frac{1}{2}}\left[(AX')^*AX'\right]^{\frac{1}{2}}} = \frac{X^*A^*AX'}{(X^*A^*AX)^{\frac{1}{2}}(X'^*A^*AX')^{\frac{1}{2}}} = \frac{X^*X'}{(X^*X)^{\frac{1}{2}}(X'^*X')^{\frac{1}{2}}} =$$

$\cos\theta$. ∎

NOTE: A special result of this theorem is that orthogonal vectors are invariant under a unitary transformation.

THEOREM 5.8. In V_n, the inner product X^*Y is invariant under a unitary transformation of coordinates.

Proof: Let U be a unitary matrix so that $X = UW$ and $Y = UZ$. Then $X^*Y = (UW)^*UZ = W^*U^*UZ = W^*Z$ which is also an inner product. ∎

THEOREM 5.9. Similar matrices have equal determinants.

Proof: Let A and B be similar matrices. Then a nonsingular square matrix C of the same order as A and B such that $C^{-1}AC = B$ exists. Then $|B| = |C^{-1}|\,|A|\,|C| = |C^{-1}|\,|C|\,|A| = |C^{-1}C|\,|A| = |I|\,|A| = |A|$. ∎

THEOREM 5.10. Similar matrices have equal eigenvalues.

Proof: Let A and B be similar matrices. Then a nonsingular square matrix C of the same order as A and B such that $C^{-1}AC = B$ exists. Then, $|A - \lambda I| = |C^{-1}(A - \lambda I)C| = |C^{-1}AC - \lambda C^{-1}IC| = |B - \lambda I|$. ∎

As stated previously, it would be impossible to describe all transformations thoroughly. However, a deeper study of one transformation is practical and perhaps desirable. Such a study of orthogonal transformations is presented in Chapter VI.

Chapter VI

A STUDY OF ORTHOGONAL TRANSFORMATIONS

One of the transformations which is often discussed in texts on matrix theory is the orthogonal transformation. Much of the material which is presented in this chapter may be found in [7]. In this presentation only real, 3 x 3 matrices will be considered.

Case 6.1. Euler's Theorem.

Euler implied in his theorem that, given the initial and final positions of a rigid body any one of whose points takes up the position from which it started, then it would have been possible to reach the final position by some one rotation about one fixed axis.

DEFINITION 6.1. A <u>rigid body</u> is composed of any number of points whose separations remain unchanged, not only after a displacement of the body, but at all times during the process of taking the body from its initial to its final position.

NOTE: Excluded will be the case of a thin rod, in which all of the points are collinear.

Let P (Figure 6.1) be any general point of a rigid body with coordinates x,y,z referred to a fixed system of rectangular axes with origin at O. Let the rigid body be displaced so that P moves to P' (coordinates x', y', z') the point at the origin (0,0,0) remaining fixed. Let the change in the coordinates of P as a result of the displacement be given by $p' = f(p)$ where p, p' are the column vectors of the coordinates of P and P' respectively. Since the point at the origin remains fixed $f(0) = 0$.

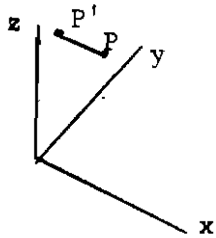


Figure 6.1

THEOREM 6.1. The separation $|x - y|$ of two vectors x and y is preserved if and only if the transformation is orthogonal.

Proof: That an orthogonal transformation preserves separation has already been proved in Theorem 5.5. It is now necessary to show that if $x' = f(x)$ describes a transformation of the total space that preserves separation and if $f(0) = 0$, then $f(x) = Ax$ where A is orthogonal.

Since the transformation preserves separation, $|f(x) - f(y)| = |x - y|$. But this is true for all y, in particular $y = 0$. Hence $|f(x) - f(0)| = |x|$. But $f(0) = 0$ and hence $|f(x)| = |x|$. That is, $x'^T x' = x^T x$. Again, if $y' = f(y)$, the preservation of separation implies that $(x' - y')^T(x' - y') = (x - y)^T(x - y)$. Thus putting $x'^T x' = x^T x$ and $y'^T y' = y^T y$ the result is $y'^T x' = y^T x$ and the transformation $f(x)$ preserves inner products. In particular, if e_i is the i^{th} column of the unit matrix and e_i' is its transform, then $e_i'^T e_j' = e_i^T e_j = 1$ when $i = j$ and $e_i'^T e_j' = e_i^T e_j = 0$ when $i \neq j$.

The vectors e_i' $(i = 1,2,\ldots,n)$ thus form an orthonormal set and a matrix A whose i^{th} column $A_{*i} = e_i'$ is orthogonal. Moreover, $e_i' = Ae_i$.

Suppose that $u_1, u_2, \ldots, u_n$ form an orthonormal set, then they are linearly independent and it is possible to find scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that any vector V can be expressed $V = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n$. Premultiplying both sides by u_i^T, then $u_i^T V = \alpha_i$. Hence $V = u_1^T V \cdot u_1 + u_2^T V \cdot V + \ldots + u_n^T V \cdot u_n$. Putting $V = f(x) = x'$ and $U_i = e_i'$, the result is $f(x) = e_1'^T x' \cdot e_1' + e_2'^T x' \cdot e_2' + \ldots + e_n'^T x' \cdot e_n'$. But inner products are preserved by the transformation, hence $e_i'^T x' = e_i^T x$. Hence, $f(x) = e_1^T x \cdot e_1' + e_2^T x \cdot e_2' + \ldots + e_n^T x \cdot e_n'$. But $e_i' = Ae_i$, thus $f(x) = e_1^T x \cdot Ae_1 + e_2^T x \cdot Ae_2 + \ldots + e_n^T x \cdot Ae_n = x_1 A_{*1} + x_2 A_{*2} + \ldots + x_n A_{*n} = Ax$. ∎

Thus the displacement of a body such that one point remains fixed and separations are preserved is algebraically represented by the linear transformation $p' = Ap$ where A is orthogonal.

Separation must also be preserved at all stages of the continuous process of taking P to P'. Consider the continuous displacement of P' back to P in the opposite sense. The elements of A must continuously approach those of an orthogonal matrix representing no displacement. This matrix is the identity matrix I, for if $A = I$, then $p' = Ip = p$. But an orthogonal matrix has a determinant which is either +1 or -1 (Theorem 5.1) while $|I| = +1$. It is impossible that A can continuously approach I if its determinant is -1. Therefore, all displacements of a rigid body with the origin fixed are represented by <u>proper</u> orthogonal transformations. (Definition 5.2).

Since the determinant of a matrix is the product of its characteristic roots, it is possible for the characteristic roots of A to be (1,-1,-1), (1,1,1), (-1,-1,-1), or (1,1,-1). Considering only the case $|A| = +1$, the latter two possibilities may be eliminated. Since A is a real orthogonal matrix and all real orthogonal matrices are unitary (Theorem 3.2) and, in turn, all unitary matrices are normal (Theorem 3.1), then it follows that A is normal. This means that $A = PDP^*$ where P is a unitary matrix and D is a diagonal matrix which consists of the eigenvalues of A. (Theorem 3.12).
Considering the two possibilities that remain, D either looks like $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. Both of these matrices have invariant subspaces of dimension one. A subspace of dimension one is a line which passes through the origin.

Consider Figure 6.2, any two points A, B on the fixed line through O. Let P be any other point not on the line. Since the transformation leaves AP, BP invariant and A and B fixed, the locus of P is a circle normal to AB. Similarly the locus of Q is also a circle normal to AB. But PQ is invariant and hence the planes ABQ and ABP rotate around AB through the same angle. Since this argument applies to all pairs of points of the body not on AB, the transformation $p' = Ap$ induces a rotation of the body about AB.

The fixed line through the origin is known as the Axis of Rotation. Therefore, all continuous displacements of a rigid body with one point fixed can be represented by orthogonal transformations and these transformations are rotations.

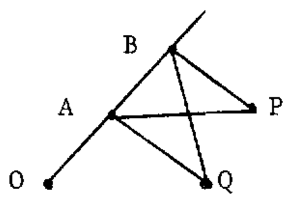


Figure 6.2

Case 6.2. The Resultant of Several Rotations.

Let the rotations be represented by proper orthogonal matrices $R_1, R_2, \ldots R_n$ and let them take place in that order. Let the coordinate vector of a point P be p. Then after the first rotation $p_1 = R_1 p$ and after the second $p_2 = R_2 p_1 = R_2 R_1 p$. The resultant is then $p_n = R_n \ldots R_2 R_1 p$.

The displacement of P from p to p_n is equivalent to a single rotation (the end result is the same) whose representative matrix R is given by $R = R_n \ldots R_2 R_1$. This is the geometrical equivalent of the property of orthogonal matrices that the product of any number of them is orthogonal.

Orthogonal matrices are, in general, non-commutative in multiplication. The right-hand matrix factor corresponds to the first rotation and so on. If $p' = Rp$ then $p = R^{-1}p' = R^T p'$ since R is orthogonal. The matrix R^T thus represents the inverse rotation to R.

Case 6.3. Given two arbitrary, fixed, distinct points and two arbitrary, fixed, distinct axes, is it possible to find a transformation which is a composition of two rotations, one about each of the lines, so that P is displaced to P'?

Take as three points the origin O and any two distinct points A and B a unit distance from the origin. If a and b are the coordinate vectors of the points A and B respectively, then they must be unit vectors. (Figure 6.3).

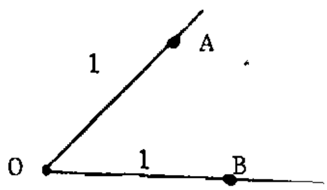


Figure 6.3

If all displacements are possible, then select any point a unit distance from O and move it to coincide with an arbitrarily chosen point also a unit distance from the origin.

Let OA and OB (Figure 6.3) be two fixed axes of rotation. The question arises: Is it possible to achieve all possible displacements by one, properly chosen, rotation about OA followed by one rotation about OB? The answer is no. Consider a point of the body initially coincident with A, then the rotation about OA leaves the point fixed at A, and the rotation about OB cannot then move the point to B. If the required rotations about OA and OB are represented by proper orthogonal matrices A and B, then for the first rotation $Aa = a$ and the second rotation gives $Ba = a'$. Is it possible that a' and b are the same unit vector? If so, then $b^T a' = 1$ that is, $b^T a' = b^T Ba = 1$. Since $Bb = b$, then $b^T B^T = b^T$. Hence, $b^T B^{-1} = b^T$ (since B is orthogonal). So, $b^T = b^T B$ and $b^T a = b^T Ba = 1$ which is true if and only if a and b are the same unit vector, but this is contrary to the hypothesis that A and B are distinct points.

Case 6.4. Rotations about Three Fixed Axes.

Consider one rotation about each of three fixed axes in a given sequence. Let the chosen axes be OA, OB, and OC represented respectively by the unit vectors a, b, c and let the rotation take place around these axes in the sequence of the letters. (Figure 6.4)

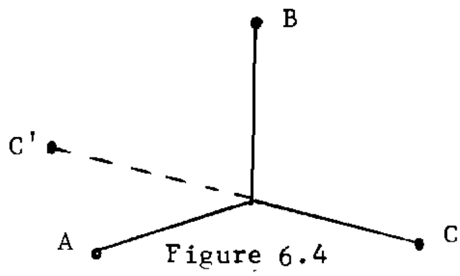


Figure 6.4

If all displacements are possible, then OB must necessarily be perpendicular to both OA and OC. To show this, it is necessary to try to take a point initially at A to C. After the first rotation A, $Aa = a$. After the second rotation B, $Ba = a'$ and after the third rotation C, $Ca' = c$ and $a' = C^T c = c$. Hence $b^T c = b^T a' = b^T Ba = b^T a$ and OB is equally inclined to OA and OC.

It is also necessary to be able to take the point initially at A to C' where C' lies on CO produced a unit length beyond O. The coordinate vector of C' is -c. By the same argument, $b^T c = -b^T a$ and $2b^T a = 0$ ($b^T c = b^T a = -b^T a = 0$). OB is thus perpendicular to OA and also to OC ($\cos\theta = \frac{b^T a}{1} = \frac{0}{1} = 0$ so $\theta = 90^\circ$; similarly for $\cos\phi = \frac{b^T a}{1} = \frac{0}{1} = 0$ so $\phi = 90^\circ$).

To show that this perpendicularity condition is sufficient, it becomes necessary to show that rotations about OA, OB, and OC can take two points initially coincident with A and B into any general position consistent with their remaining the same (unit) distance from the origin and the same distance apart. Starting with two points P, Q arbitrarily placed, but subject to the requirement that OP and OQ are of unit length and such that PQ = AB, it is possible by reversing the rotations and sequence of rotation, to bring P and Q into coincidence with A and B respectively. (Figure 6.5).

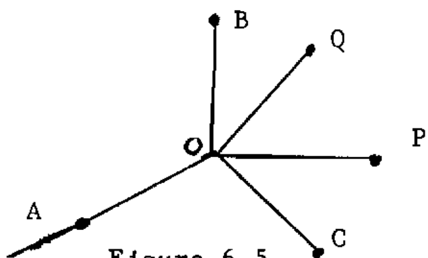


Figure 6.5

Plane AOC divides the space into two parts. Consider any point of the rigid body, say P; it either lies in the plane AOC, or a rotation around OC which lies in the plane will carry it from one half-space to the other taking it through the plane in doing so. When it lies in AOC, OP will be perpendicular to OB which is normal to the plane. A rotation about OB will then carry P into coincidence with A; and a third rotation about OA will then carry Q into coincidence with B without disturbing the coincidence of P and A. Carrying out these rotations in reversed order, it is possible to take two points initially in coincidence with A and B into a general position P, Q.

Case 6.5. The Forms of Orthogonal Matrices that Represent Rotations about a Fixed Coordinate Axis.

It is necessary to first derive the forms of the orthogonal matrices R_x, R_y, and R_z that represent rotations about the (fixed) coordinate axes O_x, O_y, and O_z respectively.

Consider first a rotation θ_x about O_x (Figure 6.6a) under which a point P, coordinates x,y,z is carried to P', coordinates x',y',z'. A projection of the space points to the yz-plane is as shown in Figure 6.6b.

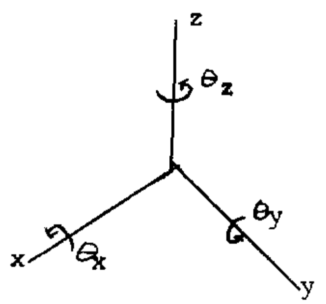


Figure 6.6a

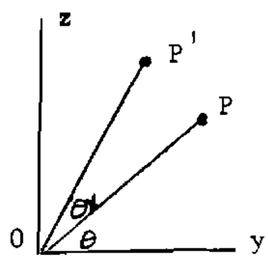


Figure 6.6b

If the projection OP makes an angle θ with O_y, and the projected length OP (=OP') is s, then $x' = x$; $y' = s\cos(\theta + \theta_x)$, $y = s\cos\theta$; $z' = s\sin(\theta + \theta_x)$, $z = s\sin\theta$ where $y' = s\cos\theta\cos\theta_x - s\sin\theta\sin\theta_x = y\cos\theta_x - z\sin\theta_x$ and $z' = s\sin\theta\cos\theta_x + s\cos\theta\sin\theta_x = z\cos\theta_x + y\sin\theta_x$.

The above relationships may be expressed as:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The matrix is proper orthogonal and may be expresses as R_x.

Similarly R_y (rotations about O_y) = $\begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix}$ when $x' = s\sin(\theta + \theta_y)$, $x = s\sin\theta$; $y' = y$; and $z' = s\cos(\theta + \theta_y)$, $z = s\cos\theta$.

Finally, R_z (rotations about O_z) = $\begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}$ when $x' = s\cos(\theta + \theta_z)$, $x = s\cos\theta$; $y' = s\sin(\theta + \theta_z)$, $y = s\sin\theta$; and $z' = z$.

The three rotations can be combined in six possible ways corresponding to the six possible rotation sequences. (Table 6.1).

SEQUENCE OF ROTATION			RESULTANT MATRIX
FIRST	SECOND	THIRD	
O_x	O_y	O_z	$R_zR_yR_x$
O_x	O_z	O_y	$R_yR_zR_x$
O_y	O_x	O_z	$R_zR_xR_y$
O_y	O_z	O_x	$R_xR_zR_y$
O_z	O_x	O_y	$R_yR_xR_z$
O_z	O_y	O_x	$R_xR_yR_z$

Table 6.1

The intermediate rotation takes place about an axis perpendicular to the other two; and any of the six permutations of the three matrices can be selected to simulate all rotations of a rigid body by rotations about three fixed axes.

Case 6.6. A Rotation of Coordinate Axes (Transformation of Coordinates).

Consider the problem of a fixed body whose coordinates are referred to a rotated system of axes. Consider a rotation (of a rigid body) represented by an orthogonal matrix R.

Let P, Q, R be three points of the body initially unit distances along the coordinate axes (Figure 6.7). Their initial coordinate vectors will be e_1, e_2, e_3; the columns of the unit matrix. After the rotation the points will move to P', Q', R' where OP', OQ', OR' are all unity, the coordinate unit vectors p', q', r' defining three mutually perpendicular directions Ox', Oy', Oz'. Then, $p' = Re_1 = R_{*1}$, $q' = Re_2 = R_{*2}$, $r' = Re_3 = R_{*3}$. The coordinate vectors p', q', r' are the columns of R and the elements of R are the rotated coordinates of points initially situated at points unit distances along the axes. Since p', q', r' or (R_{*1}, R_{*2}, R_{*3}) are unit vectors, they are the direction cosines of Ox', Oy', Oz' respectively with respect to Ox, Oy, Oz. (Figure 6.7)

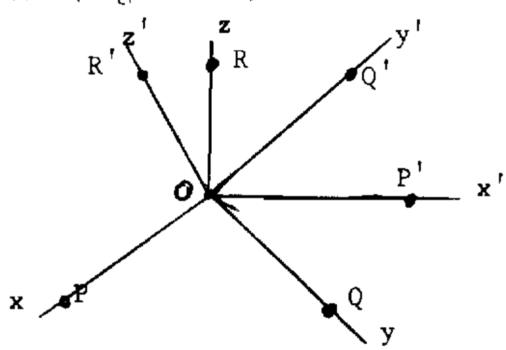


Figure 6.7

Regard the three mutually perpendicular lines Ox', Oy', Oz' as the axes of a second coordinate system with origin at O. Let a be the coordinate vector of a point A fixed with respect to Ox, Oy, Oz and let a' be its coordinate vector with respect to Ox'. Oy', Oz'.

A could be a fixed point (which it is) relative to which Ox', Oy', Oz' have been rotated from their initial positions of coincidence with Ox, Oy, Oz or A could have formed part of the rigid body containing Ox', Oy', Oz'. In this latter case the coordinate vector a' of A relative to Ox', Oy', Oz' would have remained fixed, but its coordinate vector a, relative to Ox, Oy, Oz would have changed as the body was rotated into its final position. The problem is reduced from one of transformation of coordinates to one of rigid body rotation. It is now possible to write $a = Ra'$ or $a' = R^T a$ which tells how its coordinates change with the axes as A is held fixed. Thus, if the rotation of a new set of axes with respect to an old set is represented by an orthogonal matrix R (columns of R are the direction cosines of the new axes with respect to the old) then the new and old coordinate vectors of a fixed point are related by the equation $a' = R^T a$.

Case 6.7. A Transformation of Orthogonal Matrices.

An orthogonal matrix A represents a certain rotation of a rigid body, but with respect to a chosen rectangular system of coordinates. In the equation $p_2 = Ap_1$ the coordinate vectors p_1 and p_2 together with A are referred to a given system of coordinates. It is possible to say that a point is displaced from P_1 to P_2 such that P_1 is transformed to P_2 or it is possible to refer coordinates to a new (accented) system and the <u>same</u> displacement would be represented as a transformation of P_1' to P_2' and the <u>same</u> rotation would be effected by an orthogonal matrix A' which would not usually have the same form as A. It is desirable to find the relationship between A' and A when the accented coordinate system is obtained by a rotation R about the origin of the unaccented system. (Figure 6.8).

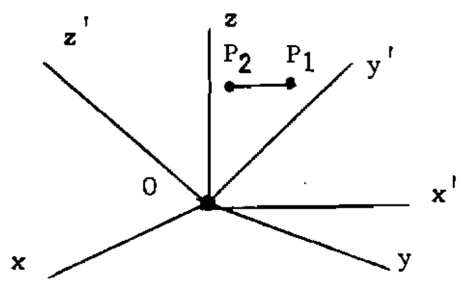


Figure 6.8

With reference to Ox, Oy, Oz a rotation represented by the matrix A carries a point P_1 of a rigid body to P_2. With reference to Ox', Oy', Oz' the same rotation is represented by A'. So $p_2 = Ap_1$ and $p_2' = A'p_1'$ where p_2 and p_2' are coordinate vectors representing the <u>same</u> points P_2 and P_1 respectively but with reference to two systems of coordinates. Thus $p_1' = R^T p_1$ and $p_2' = R^T p_2$ (see Case 6.5), so $R^T p_2 = A'R^T p_1$; that is, $p_2 = RA'R^T p_1$ and $p_2 - p_2 = RA'R^T p_1 - Ap_1$, so $0 = (RA'R^T - A)p_1$. Since this is to hold for all points, p_1 is arbitrary and $A = RA'R^T$ or $A' = R^T AR$. A very important property of the transformation of A to A' is that the eigenvalues of A' are the same as those of A. The proof of this is as follows: $|\lambda I - A| = |R^T|\,|\lambda I - A|\,|R| = |\lambda R^T R - R^T AR| = |\lambda I - A'|$.

Case 6.7. Improper Orthogonal Matrices - Transformations that Represent Reflections.

Let E be a real 3 x 3 diagonal matrix such that its diagonal elements are ± 1 with an odd number of negative signs. Hence $E = -1$

and the inverse of an E-type matrix is also an E-type matrix. Every E-type matrix is improper orthogonal.

Let A be an improper orthogonal matrix, then AE is proper orthogonal for $|AE| = |A|\,|E| = -1 \cdot -1 = +1$. If $B = AE$, then $A = BE^{-1}$ and any improper orthogonal matrix may be factorized into a proper orthogonal matrix and an E-type matrix. Moreover, the kind of E-type matrix is arbitrarily chosen. An improper orthogonal transformation is then equivalent to a proper orthogonal matrix (rotation) preceded or followed by a type-E transformation. Consider the E-type transformation $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, so $x' = x$, $y' = y$, and $z' = -z$.

This is a reflection in the xy-plane. The three E-type matrices with one negative sign thus represent reflections in the coordinate planes; when all three elements are negative, the transformation is a reflection in the origin: that is, a point is translated into the opposite octant. Thus every improper orthogonal matrix is equivalent to a rotation preceded or followed by a reflection of one of the given types.

Consider any plane through the origin normal to the unit vector n. Take this normal as the x'-axis of a new (accented) coordinate system. If R represents the rotation of this new system with respect to the old, then n will be the first column of R'.

With respect to this new system, the improper orthogonal matrix A will become $A' = R^T AR$; A' is improper and may be factorized into a rotation B and a reflection in the plane (the $y'z'$-plane of the new coordinate system) the normal to which is n. Thus $A' = B'E'$ where $E' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $A = RA'R^T = RB'E'R^T$. Transforming back to the unaccented system putting $B' = R^T BR$, then $A = RR^T BRE'R^T = BRE'R^T$. But $RE'R^T$ is the matrix representing the reflection in the given plane with reference to the unaccented system. Hence an improper matrix can be factorized into a rotation and a reflection in any plane through the origin.

To summarize Chapter VI, an orthogonal matrix represents a transformation that is either a rotation or a reflection.

Chapter VII

CONCLUSION

The main objective of this thesis has been to expose the relationships between specific classes of matrices which are often alluded to, but seldom discussed in detail in books on matrix theory. A second objective has been to present the reader with a discussion of what certain transformations accomplish.

There are many possibilities for further study in this area of matrix theory. There are other special classes of matrices which were not defined in this paper. How these matrices are related to each other and to those presented in this thesis would be an interesting area of study.

The transformations that matrices represent lend themselves very readily to a more thorough research. A discussion of what orthogonal transformations represent geometrically was presented in this paper. Other classes of matrices also represent particular transformations. What these transformations do geometrically and how these transformations compare to each other would also be excellent areas for study.

BIBLIOGRAPHY

1. Finkbeiner, Daniel T. Introduction to Matrices and Linear Transformations. San Francisco: W. H. Freeman, 1960.

2. Hohn, Franz E. Elementary Matrix Algebra. New York: Macmillan, 1958.

3. Marcus, Marvin and Henryk Minc. A Survey of Matrix Theory and Matrix Inequalities. Boston: Allyn and Bacon, 1964.

4. Pettofrezzo, Anthony J. Matrices and Transformations. Englewood Cliffs, N. J.: Prentice-Hall, 1966.

5. Schwerdtfeger, Hans. Introduction to Linear Algebra and the Theory of Matrices. Holland: Gronigen, 1961.

6. Smiley, Malcolm F. Algebra of Matrices. Boston: Allyn and Bacon, 1965.

7. Thompson, E. H. An Introduction to the Algebra of Matrices with Some Applications. Toronto: University of Toronto Press, 1969.

8. Thrall, Robert M. and Leonard Tornheim. Vector Spaces and Matrices. New York: Wiley, 1957.

9. Zelinsky, Daniel. A First Course in Linear Algebra. New York: Academic Press, 1968.